UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/685,586	10/16/2003	Daben Liu	BBNT-PO1-086	5146

28120 7590 09/28/2007
ROPES & GRAY LLP
PATENT DOCKETING 39/41
ONE INTERNATIONAL PLACE
BOSTON, MA 02110-2624

EXAMINER	
SIEDLER, DOROTHY S	
ART UNIT	**PAPER NUMBER**
2626	
MAIL DATE	**DELIVERY MODE**
09/28/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)
	10/685,586	LIU ET AL.
	Examiner	Art Unit
	Dorothy Sarah Siedler	2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on 06 July 2007.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 1-31 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1-31 is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on 16 October 2003 is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date 4-9-07,5-21-07,6-2-07,9-18-07.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

This office action is in response to the amendment filled July 6, 2007. Claims 1-31 are pending, with independent claims 1,11,18, 23, and 30 amended.

Response to Amendment

Applicant has successfully amended independent claims 1,11,18 and 30. Therefore the 35 U.S.C. §101 rejection of claims 1-22 and 30-31 is withdrawn.

Response to Arguments

Applicant's arguments with respect to claims 1-31 have been considered but are moot in view of the new ground(s) of rejection.

Specification

The disclosure is objected to because of the following informalities: The specification discloses that speaker change is detected by calculating a generalized likelihood value (λ) for two cepstral vectors. As stated on page 16, "If λ is above a predetermined threshold, the two vectors are considered to be similar to one another, and are assumed to be from the same speaker. Otherwise, the two vectors are dissimilar to one another, and the boundary point corresponding to the two vectors is defined as a speaker change. " However, this is contrary to the standard hypothesis from the generalized ratio test, as indicated in ***Fancourt*** ("On the use of Neural Networks in the Generalized Likelihood Ratio Test for Detecting Abrupt Changes in Sign

als" IEEE 2000). The optimized likelihood creates a decision function, which indicates a change when the value *exceeds* a threshold.

Appropriate correction is required.

Claim Rejections - 35 USC § 112

The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

Claim 6 is rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claim 6 recites, "wherein a speaker change is detected when the generalized likelihood ratio test produces a value less than a preset threshold", however this is ambiguous, since it is contrary to the standard hypothesis from the generalized ratio test. As noted previously, the generalized likelihood ratio test is a test for two hypotheses. The optimized likelihood creates a decision function, which indicates a change when the value *exceeds* a threshold, as indicated in ***Fancourt***.

The examiner interprets claim 6 in terms of ***Fancourt***, i.e. a speaker change is detected if the optimized likelihood value exceeds a specific threshold. This interpretation is used throughout the remainder of this office action.

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 1,7,11,14 and 30 are rejected under 35 U.S.C. 102(b) as being anticipated by ***Kubala*** ("Integrated Technologies for Indexing Spoken Language" ACM 2000).

As per claim 1, ***Kubala*** discloses a method for detecting speaker changes in an input audio stream comprising:

Segmenting the input audio stream into predetermined length intervals (page 53, first paragraph, *the speech is input as frames*);

Decoding the intervals to produce a set of phones corresponding to each of the intervals (page 53, *phone class recognition is performed on each frame. Therefore it is inherent that a set of phones was decoded for each frame*);

Generating a similarity measurement based on a first portion of the audio stream that is within one of the intervals and that occurs prior to a boundary between adjacent phones in one of the intervals and a second portion of the audio stream that is within the one of the intervals and that occurs after the boundary (page 53, second paragraph,

speaker change is hypothesized at every phone boundary using a form of a likelihood ratio test (similarity score); and

Detecting speaker changes based on the similarity measurement (page 53, second paragraph, *speaker change is hypothesized at every phone boundary using a form of a likelihood ratio test (similarity score)*; and

Outputting an indication of the detected speaker changes (page 49, *the system outputs a summary which indicates the location of speaker changes within the recording*).

As per claim 7, ***Kubala*** discloses the method of claim 1, wherein the decoded set of phones is selected from a simplified corpus of phone classes (page 53, first paragraph, *speech is labeled using speech and non-speech models*).

As per claim 11, ***Kubala*** discloses a device for detecting speaker changes in an audio signal, the device comprising:

A processor (page 52, second column, first paragraph, *the Pentium II processor*);

A memory (page 52, *the system is run on a computer using a Pentium II processor, therefore it is inherent that there are instruction stored in memory*) containing instructions that when executed by the processor cause the processor to:

Segment the audio signal into predetermined length intervals (page 53, first paragraph, *the speech is input as frames*),

Decode the intervals to produce a set of phones corresponding to each of the intervals (page 53, *phone class recognition is performed on each frame. Therefore it is inherent that a set of phones was decoded for each frame*),

Generate a similarity measurement based on a first portion of the audio signal that occurs prior to a boundary between phones in one of the sets of phones of an interval and a second portion of the audio signal that occurs after the boundary (page 53, second paragraph, *speaker change is hypothesized at every phone boundary using a form of a likelihood ratio test (similarity score)*, and

Detect speaker changes based on the similarity measurement boundary (page 53, second paragraph, *speaker change is hypothesized at every phone boundary using a form of a likelihood ratio test (similarity score)*, and

Store an indication of the detected speaker changes (page 49 and 52, and Figure 5, *the structural summarization, including the locations of speakers in a recording, are stored in an XML file in the Indexer subsystem*).

As per claim 14, this claim contains limitations that are similar to the limitations cited in claim 7, and is rejected for similar reasons.

As per claim 30, ***Kubala*** discloses a device comprising:

Means for segmenting the input audio stream into predetermined length intervals (page 53, first paragraph, *the speech is input as frames*);

Means for decoding the intervals to produce a set of phones corresponding to each of the intervals (page 53, *phone class recognition is performed on each frame. Therefore it is inherent that a set of phones was decoded for each frame*);

Means for generating a similarity measurement based on audio within one of the intervals that is prior to a boundary between adjacent phones and based on audio within the one of the intervals that is after the boundary (page 53, second paragraph, *speaker change is hypothesized at every phone boundary using a form of a likelihood ratio test (similarity score)*;

Means for detecting speaker changes based on the similarity measurement (page 53, second paragraph, *speaker change is hypothesized at every phone boundary using a form of a likelihood ratio test (similarity score)*; and

Means for outputting the detected speaker changes (page 49 and 52, and Figure 5, *the structural summarization, including the locations of speakers in a recording, are stored in an XML file in the Indexer subsystem*).

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 2 and 12 are rejected under 35 U.S.C. 103(a) as being unpatentable over ***Kubala*** in view of ***Beigi*** ("A Distance Measure Between Collection of Distributions and it's Application to Speaker Recognition" IEEE 1998).

As per claim 2, ***Kubala*** discloses the method of claim 1, however ***Kubala*** does not disclose wherein the predetermined length intervals are approximately thirty seconds in length. ***Beigi*** discloses the use of intervals that are approximately thirty seconds long in a speech recognition system (page 756, Section 4 Results, *30 seconds of speech is used for training*). ***Beigi*** discloses a system that models speech for different speakers as two sets of statistical distributions. A meaningful distance measure between the two distributions is calculated, which can then be used for speaker classification, speech segmentation and speaker verification. ***Beigi*** also uses thirty seconds of speech data as enrollment data. Therefore, the examiner argues that it is old and well known to segment audio into predetermined intervals approximately thirty seconds long.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use predetermined length intervals approximately thirty seconds in length in ***Kubala***, since an interval of that length would provide robust data for training a speaker segmentation model.

As per claim 12, this claim contains limitations that are similar to the limitations cited in claim 2, and is rejected for similar reasons.

Claims 4-6,8-10, 15-18, 21 and 22 are rejected under 35 U.S.C. 103(a) as being unpatentable over ***Kubala*** in view of ***Liu*** ("Fast Speaker Change Detection for Broadcast News Transcription and Indexing" 1999).

As per claim 4, ***Kubala*** discloses the method of claim 1, however ***Kubala*** does not disclose wherein generating a similarity measurement includes: calculating cepstral vectors for the audio stream prior to the boundary and after the boundary, and comparing the cepstral vectors. ***Liu*** discloses a system for speaker change detection that calculates cepstral vectors for the audio stream prior to the boundary and after the boundary, and compares the cepstral vectors (Section 4 Speaker Change Detection, subsection Distance Measure Criterion, *cepstral vectors are used in the distance*

measure (similarity measure)). In addition, cepstral vectors are one of many vector types used to represent speech features in speech processing tasks.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to calculate cepstral feature vectors for the audio stream prior to and after the boundary in ***Kubala***, since one of ordinary skill in the art has good reason to pursue the options within his/her technical grasp to achieve the result of calculating robust and reliable feature vectors.

As per claim 5, ***Kubala*** in view of ***Liu*** discloses the method of claim 4, however ***Kubala*** does not disclose wherein the cepstral vectors are compared using a generalized likelihood ratio test. ***Kubala*** does disclose calculating a similarity measure using the Likelihood ratio test. In addition, ***Liu*** discloses wherein the cepstral vectors are compared using a generalized likelihood ratio test (Section 4 Speaker Change Detection, subsection Distance Measure Criterion). Cepstral vectors are one of many vector types used to represent speech features in speech processing tasks.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to compare the cepstral vectors using a generalized likelihood ratio test in ***Kubala***, since one of ordinary skill in the art has good reason to pursue the options within his/her technical grasp to achieve the result of calculating robust and reliable feature vectors.

As per claim 6, ***Kubala*** in view of ***Liu*** discloses the method of claim 5, and ***Liu*** further discloses wherein a speaker change is detected when the generalized likelihood ratio test produces a value that exceeds a preset threshold (Section 4 Speaker Change Detection, subsection the critical region). The generalized likelihood ratio test is a standard method used to determine a change in signals. An optimized likelihood forms a decision function that is compared to a threshold, and indicates a change point when it exceeds the threshold.

Therefore it would have been obvious to one of ordinary in the art at the time of the invention to use the generalized likelihood ratio test to indicate a change when the value exceeds a threshold in ***Kubala***, since one of ordinary skill in the art has good reason to pursue the options within his or her technical grasp to make a reliable decision as to when a change in speaker has occurred in audio data.

As per claim 8, ***Kubala*** discloses the method of claim 7, however ***Kubala*** does not disclose wherein the simplified corpus of phone classes includes a phone class for vowels and nasals, a phone class for fricatives, and a phone class for obstruents. ***Liu*** discloses wherein the simplified corpus of phone classes includes a phone class for vowels, nasals, fricatives and obstruents (Section 3 Phone-Class Decode, Figure 1).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have corpus of phone classes that include a phone class for vowels and nasals, a phone class for fricatives, and a phone class for obstruents in ***Kubala***, since vowels and nasals are similar in that they both have pitch and high energy, and

can therefore be combined to significantly speed up processing, as indicated in ***Liu*** (section 3).

As per claim 9, ***Kubala*** in view of ***Liu*** discloses the method of claim 8, and ***Kubala*** further discloses wherein the simplified corpus of phone classes further includes a phone class for music, laughter, breath and lip-smack (page 53, first paragraph). However, ***Kubala*** does not explicitly disclose wherein the simplified corpus of phone classes further includes a phone class for silence. ***Liu*** discloses wherein the simplified corpus of phone classes further includes a phone class for silence (section 3).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have a phone class for silence in ***Kubala***, since non-speech events posses valuable information about speaker changes, and can be used to more accurately determine speaker changes, as indicated in ***Liu*** (section 3, first paragraph).

As per claim 10, ***Kubala*** discloses the method of claim 7, however ***Kubala*** does not disclose wherein the simplified corpus of phone classes includes approximately seven phone classes. ***Liu*** discloses wherein the simplified corpus of phone classes includes approximately seven phone classes (section 3).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have a corpus of phone classes including approximately seven phone classes in ***Kubala***, since it reduces the number of active nodes during decoding, and thus speeds up processing, as indicated in ***Liu*** (section 3, last paragraph).

As per claim 15, this claim has limitations similar to claim 8, and is therefore rejected for similar reasons.

As per claim 16, this claim has limitations similar to claim 9, and is therefore rejected for similar reasons.

As per claim 17, this claim has limitations similar to claim 10, and is therefore rejected for similar reasons.

As per claim 18, ***Kubala*** discloses a device for detecting speaker changes in an audio signal, the device comprising:

A segmentation component configured to segment the audio signal into predetermined length intervals (page 53, first paragraph, *the speech is input as frames*);

A phone classification decode component configured to decode the intervals to produce a set of phone classes corresponding to each of the intervals (page 53, *phone class recognition is performed on each frame. Therefore it is inherent that a set of phones was decoded for each frame*); and

A speaker change detection component configured to detect locations of speaker changes in the audio signal based on a similarity value calculated over a first portion of the audio signal that occurs prior to a boundary between phone classes in one of the intervals and a second portion of the audio signal that occurs after the boundary in the one of the intervals boundary (page 53, second paragraph, *speaker change is hypothesized at every phone boundary using a form of a likelihood ratio test (similarity score)*;

Wherein an indication of the detected locations of speaker changes are output from the device (page 49, *the system outputs a summary which indicates the location of speaker changes within the recording*).

However, ***Kubala*** does not disclose wherein a number of possible phone classes being approximately seven. ***Liu*** discloses wherein the simplified corpus of phone classes includes approximately seven phone classes (section 3).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have a corpus of phone classes including approximately seven phone classes in ***Kubala***, since it reduces the number of active nodes during decoding, and thus speeds up processing, as indicated in ***Liu*** (section 3, last paragraph).

As per claim 21, this claim contains limitations similar to those on claim 8, and is therefore rejected for similar reasons.

As per claim 22, this claim contains limitations similar to those on claim 9, and is therefore rejected for similar reasons.

Claims 19 and 20 are rejected under 35 U.S.C. 103(a) as being unpatentable over ***Kubala*** in view of ***Liu*** as applied to claim 18 above, and further in view of ***Beigi***.

As per claim 19, this claim contains limitations similar to those on claims 2 and 12, and is therefore rejected for similar reasons.

As per claim 20, this claim contains limitations similar to those on claims 3 and 13, and is therefore rejected for similar reasons.

Claims 3,13,23-26 and 31 are rejected under 35 U.S.C. 103(a) as being unpatentable over ***Kubala*** in view of ***Siegler*** ("Automatic Segmentation, Classification and Clustering of Broadcast News audio" 1997).

As per claim 3, ***Kubala*** the method of claim 1, however ***Kubala*** does not disclose wherein segmenting the input audio stream includes: creating the predetermined length intervals such that portions of the intervals overlap one another. ***Siegler*** discloses a

system that divides audio data into overlapping segments (section 3. Segmentation, paragraph five, *the audio signal is analyzed using a two second sliding window, placed at every point in the audio stream*). ***Siegler*** discloses a system for the segmentation, classification and clustering of broadcast news audio, and is therefore analogous art.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to create overlapping intervals in ***Kubala***, since overlapping segments would minimize segmentation errors by ensuring that speaker boundaries are determined at word boundaries instead of in the middle of words.

As per claim 13, this claim contains limitations similar to claim 3, and is therefore rejected for similar reasons.

As per claim 23, ***Kubala*** discloses a system comprising:

An indexer configured to receive input audio data and generate a rich transcription from the audio data, the rich transcription including metadata that defines speaker changes in the audio data (page 49 and 52, and Figure 5, *the Indexer subsystem creates a structural summarization, including a transcript indicating locations of speakers in a recording*), the indexer including:

A speaker change detection component configured to detect locations of speaker changes in the audio data based on a similarity value calculated at locations in the segments that correspond to phone class boundaries (page 53, second paragraph,

speaker change is hypothesized at every phone boundary using a form of a likelihood ratio test (similarity score));

A memory system for storing the rich transcription (page 49 and 52, and Figure 5, *the structural summarization, including the locations of speakers in a recording, is stored in an XML file in the Indexer subsystem*); and

A server configured to receive requests for documents and to respond to the requests by transmitting ones of the rich transcriptions that match the requests (page 55, Information Retrieval, *information indexing and retrieval take place on the Rougn'n'Ready server*).

However, ***Kubala*** does not disclose a segmentation component configured to divide the audio data into overlapping segments of a predetermined length. ***Siegler*** discloses a system that divides audio data into overlapping segments (section 3. Segmentation, paragraph five, *the audio signal is analyzed using a two second sliding window, placed at every point in the audio stream*). ***Siegler*** discloses a system for the segmentation, classification and clustering of broadcast news audio, and is therefore analogous art.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to create overlapping intervals in ***Kubala***, since overlapping segments would minimize segmentation errors by ensuring that speaker boundaries are determined at word boundaries instead of in the middle of words.

As per claim 24, ***Kubala*** in view of ***Siegler*** disclose the system of claim 23, and ***Kubala*** further discloses wherein the indexer further includes at least one of: a speaker clustering component, a speaker identification component, a name spotting component, and a topic classification component (pages 52-55 and Figure 5).

As per claim 25, ***Kubala*** in view of ***Siegler*** disclose the system of claim 23, and ***Kubala*** discloses wherein the overlapping segments are segments of a predetermined length (page 53, first paragraph, *the speech is input as frames*).

As per claim 31, ***Kubala*** discloses the device of claim 30, however Kubala does not disclose wherein the predetermined length intervals overlap one another. ***Siegler*** discloses a system that divides audio data into overlapping segments (section 3. Segmentation, paragraph five, *the audio signal is analyzed using a two second sliding window, placed at every point in the audio stream*). ***Siegler*** discloses a system for the segmentation, classification and clustering of broadcast news audio, and is therefore analogous art.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to create overlapping intervals in ***Kubala***, since overlapping segments

would minimize segmentation errors by ensuring that speaker boundaries are determined at word boundaries instead of in the middle of words.

Claims 27-29 are rejected under 35 U.S.C. 103(a) as being unpatentable over ***Kubala*** in view of in view of ***Siegler***, as applied to claim 23 above, and further in view of ***Liu***.

As per claim 27, ***Kubala*** in view of ***Siegler*** disclose the system of claim 23, however neither disclose wherein the phone classes include a phone class for vowels and nasals, a phone class for fricatives, and a phone class for obstruents. ***Liu*** discloses wherein the simplified corpus of phone classes includes a phone class for vowels and nasals, a phone class for fricatives and a phone class for obstruents (Section 3 Phone-Class Decode, Figure 1).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have corpus of phone classes that include a phone class for vowels and nasals, a phone class for fricatives, and a phone class for obstruents in ***Kubala***, since vowels and nasals are similar in that they both have pitch and high energy, and can therefore be combined to significantly speed up processing, as indicated in ***Liu*** (section 3).

As per claim 28, ***Kubala*** in view of ***Siegler***, further in view of ***Liu*** disclose the system of claim 27, and ***Kubala*** further discloses wherein the phone classes additionally include a phone class for music, laughter, breath and lip-smack, and silence (page 53, first paragraph). However, ***Kubala*** does not explicitly disclose wherein the simplified corpus of phone classes further includes a phone class for silence. ***Liu*** discloses wherein the simplified corpus of phone classes further includes a phone class for silence (section 3).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have a phone class for silence in ***Kubala***, since non-speech events posses valuable information about speaker changes, and can be used to more accurately determine speaker changes, as indicated in ***Liu*** (section 3, first paragraph).

As per claim 29, ***Kubala*** in view of ***Siegler*** disclose the system of claim 23, however neither disclose wherein the phone classes include approximately seven phone classes. ***Liu*** discloses wherein the simplified corpus of phone classes includes approximately seven phone classes (section 3).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have a corpus of phone classes including approximately seven phone classes in ***Kubala***, since it reduces the number of active nodes during decoding, and thus speeds up processing, as indicated in ***Liu*** (section 3, last paragraph).

Claim 26 is rejected under 35 U.S.C. 103(a) as being unpatentable over ***Kubala*** in view of in view of ***Siegler***, as applied to claim 25 above, and further in view of ***Beigi***.

Kubala in view of ***Siegler*** disclose the system of claim 25, however neither disclose wherein the predetermined length is approximately thirty seconds. ***Beigi*** discloses the use of intervals that are approximately thirty seconds long in a speech recognition system (page 756, Section 4 Results, *30 seconds of speech is used for training*). ***Beigi*** discloses a system that models speech for different speakers as two sets of statistical distributions. A meaningful distance measure between the two distributions is calculated, which can then be used for speaker classification, speech segmentation and speaker verification. ***Beigi*** also uses thirty seconds of speech data as enrollment data. Therefore, the examiner argues that it is old and well known to segment audio into predetermined intervals approximately thirty seconds long.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use predetermined length intervals approximately thirty seconds in length in ***Kubala***, since an interval of that length would provide robust data for training a speaker segmentation model.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dorothy Sarah Siedler whose telephone number is 571-270-1067. The examiner can normally be reached on Mon-Thur 9:30am-5:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

DSS

TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER